ICR Workspace Working Group Teleconference Meeting Minutes
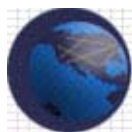
Date, Time & Location:	August 11, 2004, 2:00pm – 3:00pm EDT		
Attendees:			

First Name	Last Name	Institution
Esther	Asaki	NCI
Gary	Bader	Sloan
Craig	Beam	Moffitt
Terry	Braun	U of Iowa Holden
Thomas	Casavant	U of Iowa Holden
Leslie	Derr	NCI
Terry	Disz	U of Chicago/Argonne
Steve	Enkemann	Moffitt
Steven	Eschrich	Moffitt
David	Fenstermacher	Penn
John	Greally	Albert Einstein College of Medicine
Lynette	Grouse	NCI
David	Jewell	Dartmouth
David	Kane	NCI/SRA
Juli	Klemm	Booz Allen Hamilton/3rd Millennium
Alex	Lash	Sloan
Ted	Liefeld	Broad/MIT
Simon	Lin	Duke
James	Lyons-Weiler	Pittsburgh
Steve	Marron	UNC Lineberger
Shannon	McWeeney	Oregon Health
Arnie	Miles	Georgetown
Tom	Moloshok	Fox Chase
Rakesh	Nagarajan	Wash U
Michael	Nebozhyn	Wistar
Michael	Ochs	Fox Chase
Harold	Riethman	Wistar
John	Rux	Wistar
Carl	Schaefer	NCI
Jennifer	Shoemaker	Duke
Paul	Spellman	LBL
Philip	Stafford	TGEN
Craig	Street	Penn
Darin	Taverna	TGEN
Naveen	Vinukonda	Institute For Cancer Prevention
Kristina	Vuori	Burnham
John	Weinstein	NCI
Mathieu	Wiepert	Mayo Clinic
Cathy	Wu	Georgetown
Quing	Xaio	JPL
Liming	Yang	NCI

Integrative Cancer Research Workspace

	Yajun	Yi	Vanderbilt
	Claire	Zhu	Booz Allen Hamilton
	Julie	Zhu	Northwestern University
Introduction	**Roll-call, open meeting, review meeting goals**		
	- Summary of Special Interest Group activities		
	- Update from other Workspaces		
	- ICR face-to-face meeting		
	- General Q & A		
Review of Workspace Activities	**Summary of Special Interest Group activities**		
	- Translational Tools		
	o Terry Braun of Iowa Holden gave a Web demo of TrAPSS. The Institute for Cancer Prevention and possibly Wistar will be adopters for this program.		
	o Rembrant will be presented at the next SIG meeting. This meeting will need to be rescheduled due to the Labor Day holiday.		
	- Pathways		
	o Ajay Jain of UCSF gave a presentation of QPACA. Adopter for this application will be Oregon Health.		
	o Carl Schaefer gave a presentation of NCICB's progress on pathways data representation.		
	- Microarray Repositories		
	o Mervi Heiskanen (NCI) gave an overview of caArray data exchange. There are a number of projects within ICR that are targeting data exchange with caArray		
	o caArray user acceptance testing will be occurring at the end of this month. Will be released September 16.		
	- Genome Annotation		
	o Rakesh Nagarajan (Wash U) gave a presentation of FunctionExpress.		
	o David Kane and John Weinstein (NCI) gave a presentation of GOMiner.		
	o Wistar is the adopter for both of these projects and in general they provide complementary approaches to the annotation of microarray results.		
	- Data Analysis and Statistical Methods		
	o Steve Marron (UNC) gave a presentation of DWD.		
	o Joseph Wang (Georgetown) gave a presentation of VISDA.		
	o Wistar is also the adopter on both of these programs, which are both applications for the statistical analysis of microarray data		
	- Informatics for Proteomics		
	o David Jewell (Dartmouth) gave a presentation of Q5. The Institute for Cancer Prevention and possibly Oregon Health will adopt this application.		

	○ David is also in contact with EDRN in attempt to identify proteomics data sets that can be used for development within this SIG.

Updates from other Workspaces:	**Architecture (Presented by Simon Lin, Duke)**
	
	1. Meeting at Ohio State
	   a. Format was open discussion
	   b. Goal: To define the guidelines necessary to achieve interoperability for caBIG
	2. Major Grid projects presented:
	   a. NCICB's caGRID
	      i. Evaluated a collection of grid-enabling technologies
	      ii. Used caCORE objects to implement some scientifically interesting use cases
	   b. Ohio's Mobius
	      i. Provides middleware services and protocols for
	         1. Data and metadata management
	         2. Semantic queries across disparate data sources
	         3. Translation of data
	         4. Ad hoc creation of data warehouses
	3. Decisions that impact Silver Compatibility:
	   a. XML is the data exchange format
	      i. This should implement an object-oriented information model
	      ii. There will be support for binary attachments
	   b. SOAP is the message-exchange protocol
	   c. XML Schema will define data and message formats
	   d. Architecture will provide more guidance here as needed
	4. Decisions that impact Gold Compatibility:
	   a. Grid technology: Globus toolkit 3.2
	   b. Distributed querying: OGSA-DAI 4.0
	5. Architecture will continue to investigate:
	   a. Common query language
	   b. CDE and metadata repository
	   c. Message broker
	   d. Workflow languages/implementations
	   e. Security
	   f. XML compression
	6. Reference implementations presented:
	   a. Georgetown's PIR as a data service
	   b. Duke's RProteomics as an analysis service
	
	Controlled Vocabularies and Common Data Elements (Presented by Juli Klemm)
	
	- An External Standards Document is available that catalogs and assesses data standards relevant to clinical and basic science. This document will be reviewed at the face-to-face meeting and will be distributed to ICR Workspace this week.
	
	- The group is working on governance model for Vocabularies.
	
	Tissue Banks and Pathology Tools (Presented by Rakesh Nagarajan, Wash U)
	
	- A requirementa survey has been sent to the centers.
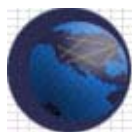
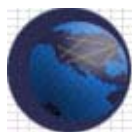|	
	- The group is currently collecting and reviewing the surveys, which have shown the range of database types at the various centers - Face-to-face meeting date was discussed, but no decision yet. Likely to schedule it just before or just after the APIII meeting in October. **Clinical Trials Management (Presented by David Fenstermacher, Penn)** - Face-to-face meeting held on July 19 & 20. - HIPAA compliance was identified as a key issue. David Fenstermacher is the contact person within caBIG for this issue. - Discussed definition of caBIG compatibility. - Additional SIG has been created– C3D SIG to focus on C3D and Oracle Clinical installation **Training (Presented by Naveen Vinkanda, Institute for Cancer Prevention)** - The Developer subcommittee is working with the Architecture Best Practices SIG to put together survey questions for developers - The Adopter subcommittee is putting together a training strategy document . **Strategic Planning (Presented by Michael Ochs)** - Relatively quiet this summer. **Data Sharing and Intellectual Capital (Presented by Terry Braun, Holden)** - A survey has been prepared to ask centers about issues around the exchange of data, specimens, etc. This would include sharing with corporate collaborators. - There were many discussions about the survey, whether it is useful or over-burdening. The group decided to go ahead and send it out.

ICR Face-to-Face Meeting	- Nearly complete participants list. 35-40 Center attendees, plus attendees from NCI and BAH
	
	- For those center who have identified attendees, those individuals should have received an email yesterday from Claire Zhu confirming their attendance and outlining how travel arrangements should be made.
	
	- Will start first thing Tuesday morning – registration beginning at 8:00 and the meeting starting at 9:00. Will end by 3:00 on Wednesday to allow participants the opportunity to travel home that evening.
	
	- Breakfast/lunch/coffee breaks will be provided on both days
	
	- Agenda: will be finalized and distributed in the next few days
	
	    o Day1: Making applications caBIG compatible
	
	        ▪ Overview and discussion of compatibility document
	
	        ▪ Introduction to model-driven architecture
	
	        ▪ Introduction to UML and object to relational mapping
	
	        ▪ Relationship between UML and CDE
	
	        ▪ How to use EVS

	▪ Harmonization
	○ Day2: ICR-specific data representations, standards and data domains
	▪ Overview of external data exchange documentation
	▪ SIG representative led discussions
General Q & A	- Each SIG seems to work on a same set of problems. How will this project avoid duplication of effort? (Steve Enkemann) ○ The goal of ICR this year is to achieve interoperability of applications across centers. It is OK to have some overlaps among SIGs, as multiple efforts are necessary to deal with different types of data. As we go, these efforts most likely will converge into a common ground. (Juli Klemm) ○ This issue has been discussed often in the Architecture Working Group. It was agreed that the short-term federated databases, and true grid should be a long-term goal. (David Fenstermacher) - Will there be training/education about the grid architecture? ○ This is somewhat of a "chicken and egg" problem. Pilots are beings designed to more closely examine the technologies that have been identified and to address outstanding issues. For now, it is important to get projects to the point where they are well-described such that they are "ready" for grid implementation. However, some retrofitting is probably inevitable (David Kane).

Action Items:	Name Responsible	Action Item	Date Due	Notes
	Juli Klemm	Distribute meeting minutes		
	Juli Klemm	Distribute draft Agenda and itinerary for the face-to-face meeting		
	Juli Klemm	Distribute the External Standards document to the Workspace		